P.B.5818 - Patentlaan 2
2280 HV Rijswijk (ZH)
☎ +31 70 340 2040
TX 31651 epo nl
FAX +31 70 340 3016

Europäisches Patentamt | European Patent Office | Office européen des brevets

Generaldirektion 1 | Directorate General 1 | Direction Générale 1

Nash, Roger William
BT Group Legal
Intellectual Property Department,
PP C5A
BT Centre
81 Newgate Str
London EC1A 7AJ
GB

Datum/Date
10/09/04

Zeichen/Ref./Réf.	Anmeldung Nr./Application No./Demande n°./Patent Nr ./Patent No./Brevet n°.
A 25570 EP	98307623.3 2416

Anmelder/Applicant/Demandeur/Patentinhaber/Proprietor/Titulaire
BRITISH TELECOMMUNICATIONS public limited company

Übersendung von/Transmission of/Envoi de

Antrag vom/Request dated/Requête du 23/08/04

- ☐ Kopien bei Akteneinsicht nach Regel 94(3) EPÜ
  Copies in the case of inspection of files pursuant to Rule 94(3) EPC
  Copies en cas d'inspection publique selon la règle 94(3) CBE
- ☐ Beglaubigung
  Certification
  Certification
- ☒ 1 Prioritätsbeleg(e)/priority document(s)/document(s) de priorité R. 94(4)
- ☐ ____ Ausfertigung(en) der Patenturkunde nach Regel 54(2) EPÜ
  Duplicate of the patent certificate pursuant to Rule 54(2) EPC
  Duplicata du certificat de brevet, selon la Règle 54(2) CBE
- ☐ Auszug aus dem Register nach Regel 92(3) EPÜ
  Extract from the register pursuant to Rule 92(3) EPC
  Extrait du registre selon la Règle 92(3) CBE
- ☐ Auskunft aus den Akten nach Regel 95 EPÜ
  Communication of information contained in the files pursuant to Rule 95 EPC
  Communication d'informations contenues dans la dossier selon la Règle 95 CBE
- ☐ Akteneinsicht nach Regel 94(2) EPÜ
  Inspection of files pursuant to Rule 94(2) EPC
  Inspection publique selon la Règle 94(2) CBE

PERON L (TEL:3632)

1158 Auskunftsstelle/Information office/Service d'information

Europäisches Patentamt | European Patent Office | Office européen des brevets

Bescheinigung | Certificate | Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. | Patent application No. | Demande de brevet n°

98307623.3

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

R C van Dijk

THIS PAGE BLANK (USPTO)

Anmeldung Nr:
Application no.: 98307623.3
Demande no:

Anmeldetag:
Date of filing: 18.09.98
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

BRITISH TELECOMMUNICATIONS public limited company
81 Newgate Street
London EC1A 7AJ
GRANDE BRETAGNE

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se referer à la description.)

Packet network

In Anspruch genommene Priorität(en) / Priority(ies) claimed /Priorité(s) revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

H04L12/00

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of filing/Etats contractants désignées lors du dépôt:

AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE

PACKET NETWORK

The present invention relates to a method of operating a packet network.

In packet network terminology, sending a packet to a single recipient is known as unicasting. Sending a packet to all possible recipients is known as broadcasting. Sending a packet to a subset of all possible recipients is referred to as multicasting.

A router in a packet network broadcasts packets by copying an input packet to all its interfaces save for the interface it arrived on. To support multicast a router must be able to copy an input packet only to those interfaces which lie on a selected path to one or more of recipients. In order to do this a router needs, for each multicast address it might have to interpret, to store an indication of the interfaces to which the packet is to be copied.

By not duplicating a packet until necessary, multicasting places less load on a packet network than would the alternative method of unicasting the same packet from the sender to each one of the selected subset of recipients.

Not all components of the present day Internet are able to process multicast packets. However, this situation is now changing rapidly and it is likely that the use of multicasting will increase significantly in the next few years.

In particular, it is to be expected that 'push' news services, Publish and Subscribe services and the like will migrate to the use of multicasting. 'Push' news services normally allow a user to select subjects of interest. One way of achieving the bandwidth savings offered by the use of multicast (whilst allowing a user to be selective as to which subjects he received) would be to allocate a multicast address to each subject. However, given the large number of possible subjects this method would result in undesirably great demands being made on the storage and processing capabilities of routers in the packet network.

According to the present invention, there is provided a method of operating a packet network to transmit a plurality of packets to respective different subsets of possible recipients, said method comprising the steps of:

for one or more portions of the network,

assigning a common local multicast address to packets are destined for different subsets of possible recipients but which share a subset of forwarding nodes within said portion.

By only requiring routers with a given portion of a network to respond to each of the possible subsets of recipients within said portion, the number of multicast addresses that must be stored by the router is reduced.

Preferably, said identification step involves:

extracting, from the payload of each packet, destination data identifying the subset of possible recipients to which the packet is to be sent; and

deriving an associated local multicast address on the basis of said destination data; and

transmitting said packet to said associated local multicast address.

This has the advantage that the invention can be implemented without alteration of network layer routing protocols.

In some embodiments, said destination data comprises subject-identifying data.

There now follows, by way of example only, a description of specific embodiments of the present invention. The description is to be read in conjunction with the accompanying drawings, in which:

Figure 1 is a schematic view of an intranet;

Figures 2A and 2B show multicast address allocations for subsections of the intranet;

Figure 3 shows the format of an application layer packet assembled in accordance with one embodiment of the present invention;

Figures 4A, 4B and 4C illustrate the building up of subject-to-IP address mappings used in the embodiment;

Figures 5A, 5B, 5C and 6 illustrate more developed subject-to-IP address mappings;

A corporate intranet (Figure 1) comprises a central site area network (CS) which interconnects three remote site area networks (S1, S2, S3). Each component of the intranet operates in accordance with the TCP/IP protocol suite and supports multicast addressing.

The central site network (CS) includes an Ethernet local area network L1 to which a sports news computer N, an organisation-level application packet forwarding computer P and a gateway router R1 are connected. The router R1 is connected to the Internet and also to a central intranet router R2.

The central intranet router R2 is connected by three communication links (e.g. optical fibres) to three respective site area ingress routers (R3, R8, R13). Each of the three ingress routers (R3, R8, R13) has a respective one of three site-level application packet forwarding computers (C1, C2, C3) attached to it.

The three site area networks (S1, S2, S3) all have a similar configuration. The second site area network S2, for example, has a central site router R9 which interconnects the site's ingress router R8 and three building routers (R10, R11, R12). Each building router (R10, R11, R12) is in turn connected to a respective one of three building area networks (L5, L6, L7) which connect a number of PCs.

The sports news computer N at the organisation's central site includes a database containing a number of news articles about various sports. These news articles are updated every few hours.

One PC (H1 to H9) in each of the building area networks (L2 to L10) is designated as a news reader PC.

In configuring the intranet, each of the site-level application packet forwarding computers (C1, C2, C3) is supplied with a table which lists Internet protocol (IP) addresses and associated sets of one or more of the on-site news reader PCs (H1 to H9). For example, site-level application packet forwarding computer C1 might be supplied with the set-to-IP address conversion table illustrated in Figure 2A. The left-hand column of that table lists the possible sets of news reader PCs (H1 to H9) and the right-hand column lists associated IP addresses. Analogous set-to-IP address conversion tables are supplied to the other two site-level forwarding computers (C2, C3).

In the next stage of the configuration process, each news reader PC is operated to request membership of the multicast groups in which it is included. The membership request is made in accordance with the Internet Group Management Protocol. For example, H1 will be operated to request membership of the uppermost three multicast groups of Figure 2A.

Similar procedures are carried out in relation the other site area networks (S1, S2).

Those skilled in the art will realise that the central site routers (R4, R9, R14) will, after this stage, include, in relation to the news application, four multicast groups.

The next stage in the configuration procedure is to supply a set-to-IP address conversion table to the organisation level packet forwarding computer P. This table is similar to that (Figure 2A) supplied to each of the site-level forwarding computers (C1, C2, C3), but the sets in this table are sets of site-level forwarding computers (C1,C2,C3) rather than sets of news reader PCs (H1 to H9). One possible example of the contents of the set-to-IP address conversion table stored in the organisation-level forwarding computer P is shown in Figure 2B.

Further, each site-level forwarding computer (C1, C2, C3) is operated to request membership of the multicast groups in which it is included. The membership request is made in accordance with the Internet Group Management Protocol. For example, site-level forwarding computer C2 will request membership of the groups identified with the first, second and fourth multicast addresses shown in Figure 2B.

Those skilled in the art will realise that this will result in central router R2 having four news application related multicast addresses in its multicast routing table.

Once this configuration has taken place, users of the news reader PCs (H1 to H9) can subscribe to news articles about their chosen sport. Each news reader PC is controlled by a news reader program which allows a user to select (using the keyboard or a Graphical User Interface for example) sports for which they would like to received news articles held in the news computer N. On such a request being made the program controls the PC to send a request packet (Figure 3 - sent using the Transmission Control Protocol) for news articles concerning the chosen sport to site-level forwarding computer (C1, C2, C3).

In the composition of the request packet, the news reader program controls the computer to provide data indicating that a file is sought 5, an indication that the files relates to a news application 10 and an identification of the subject 20 about which the news article is sought. The computer then operates to successively encapsulate the data in a TCP header 30, and an IP header 40. The TCP header 30 includes both the 'port' of the site-level computer to which request is to be sent and the 'port' of the news reader PC from which it was sent. Similarly, the IP header includes both the IP address of the site-level forwarding computer and the address of the news reader PC.

On receiving the request, the site-level application packet forwarding computer (C1, C2, C3) is controlled by a forwarding table updating program to establish which of the on-site news reader PCs (H1 to H9) are now interested in news articles about the chosen sport. If the request is not already being met then the forwarding table is updated accordingly. Those skilled in the art will be able to generate suitable table updating program.

Were news reader PC H1 to send, for example, a request for news articles concerning the subject 'pool', then the site-level forwarding computer operating under control of the program would update the forwarding table as shown in Figure 4A. It will be seen that the subject 'pool' is associated with H1's unicast IP address.

Once the table has been updated, the packet forwarding program controls the site-level application packet forwarding computer (C1, C2, C3) to check to see whether the forwarding computer (C1, C2, C3) already subscribes to the requested subject. If it does not then the site-level forwarding computer is further controlled to send a request to the organisation-level forwarding computer P. The request is of a similar format to that previously generated by the news reader PC H1 and as described above.

The organisation-level forwarding computer P is then controlled by a program similar to that described above in relation to the site-level forwarding computer to establish which set of the site-level forwarding computers (C1,C2,C3) now requires news articles concerning the selected subject to be sent. The organisation-level forwarding computer is then further controlled to update its forwarding table accordingly. Again programs to cause this operation can be easily generated by those skilled in the art.

To continue with the earlier example, in response to the request from site-level application packet forwarding computer C1 for 'pool' articles, the organisation-level computer will update its forwarding table as shown in Figure 4B. It will be seen that the forwarding table indicates that news articles concerning 'pool' are to be unicast to the IP address of site-level forwarding computer C1.

Using similar procedures to those described above, a request for articles relating to 'pool' might later be received by site-level forwarding computer C1 from news reader PC H2. The site-level packet forwarding computer (C1,C2,C3) is therefore controlled to amend its forwarding table to reflect the fact that both news reader H1 and news reader H2 should be sent articles concerning 'pool'. The site-level

forwarding computer will therefore update its forwarding table as shown in Figure 4C.

It will be seen that as two news reader PCs (H1, H2) have now requested articles relating to a subject, the IP address associated with the subject is a multicast address (in the notation generally used, IP multicast addresses begin with a number between 224 and 239 – unicast addresses begin with lower numbers).

It will also be realised by those skilled in the art that 'unsubscribe' requests could also easily be implemented.

It will be realised that after a number of requests have been received from the various news reader PCs (H1 to H9) in the intranet, the forwarding tables will associate several subjects with each multicast or unicast address.

For example, the forwarding tables of site-level forwarding computers C1, C2 and C3 might be as shown in Figures 5A, 5B and 5C respectively. At the same time, the forwarding table associated with organisation-level forwarding computer P might be as shown in Figure 6.

On receiving a news article, each of the site-level forwarding computers (C1,C2,C3) is controlled by a forwarding program to extract both the news article (Figure 3, 50) and subject 20 from the payload of the incoming TCP protocol data unit and send the news article to the IP address associated with that subject in the table. Programs executable to control the forwarding computers to behave in this manner can be generated by those skilled in the art.

To give an example, were the organisation-level computer P to receive a new news article concerning 'golf' and were the forwarding tables to be as shown in Figures 5A, 5B, 5C and 6 then the organisation-level forwarding computer P would be controlled to send the packet to IP address 229:274:1:27 (a multicast address). On receiving that packet the router R1 would, in accordance with the routing table entries made at the configuration stage, forward it to the router R2 which would, also in accordance with the routing table entries made at the configuration stage,

forward the packet further to all three of the site-level forwarding computers (C1,C2,C3) – see first row of Figure 2B. Once the packet arrives at the second site-level forwarding computer C2, that forwarding computer C2 is controlled to replace the IP address on the incoming packet with the IP address associated with the subject (golf). The third line of Figure 5B shows the associated address. The forwarding computer C2 is then controlled to forward the packet to IP address 230.81.58.3. Owing to the earlier configuration stage, the site ingress router R8 receives the packet and forwards it to the central site router R9. The central site router R9 copies the packet to its interfaces with the first and third on-site building area networks L5, L7. The building routers R10 and R12 are then operated to forward the packets to the respective news reader PCs H4 and H5.

By providing the forwarding computers with further programs, they could be rendered operable to allow efficient 'cross-posting'. 'Cross-posting' involves the sending of a message which concerns a plurality of subjects. The subject field of a forwarded packet might contain an identification of a pair of subjects, for example. It can be envisaged that the forwarding computer could operate under control of the further program to identify which set or recipients require one or both subjects. Using the techniques of the above embodiment, the selection of an appropriate multicast address would then be trivial.

Similarly, in relation to reliable multicast protocol, a further program could be used to control the forwarding computers to re-send messages only to the subset of recipients who did not acknowledge the message. Again, using the techniques of the above embodiment, the selection of an appropriate multicast address would be trivial.

It will be seen that each forwarding computer C1, C2, C3 can forward a message on one of seven possible site area trees. The site area tree used is independent of the higher level tree used in sending the packet to a set of the forwarding computers C1, C2, C3. By combining the site area trees and the higher level tree a large number of organisation area trees can, in effect, be created. Importantly, this can be done using fewer multicast addresses (twelve) than would be needed to enable multicasting directly from the organisational-level forwarding computer P.

To multicast directly from the organisational-level forwarding computer P, over five hundred multicast addresses could potentially be required.

The above embodiment only enables the transmission of news articles from the news computer N to the news reader PCs (H1 to H9). A 'publish and subscribe' service could be provided by programming the news reader PCs (H1 to H9) to be operable to unicast messages containing news articles input by their users to the news computer N. These messages could then be sent back down the hierarchy as described above.

More complex arrangements which are effective to obviate the need to send messages to the news computer at the top of the hierarchy might also be used.

It will be seen that distribution of the news articles around the intranet only requires the use of twelve multicast addresses and that this is independent of the number of subjects. The number of news subjects might be considerably greater than twelve, and in such situations the present embodiment provides significant advantages over per-subject multicast addressing.

An even more advantageous reduction occurs in the number of multicast addresses that need to be stored in the routers of the intranet. In the above embodiment, each router is only required to store details of four multicast addresses an the associated interfaces. In contrast, per-subject multicast addressing would require each router to store a number of addresses equal to the number of subjects – again in many circumstances this will be a considerably greater number.

The above advantages are amplified in embodiments that have more than two hierarchical layers in the forwarding structure. For example, an embodiment can be envisaged in which town-level and country-level forwarding computers are used.

In other embodiments many different applications might be supported. For example, by replacing the subject field of the above embodiment with a field that indicated both an application and a subject, the embodiment could support a

plurality of subject-based applications. One such application is the announcement of multicast sessions. Session announcement could include a subject identifier similar to that currently seen in Network News applications (e.g. comp.internet.ietf.conference). It will be clear how the above embodiment might be altered to allow a user to selectively receive only those announcements which concern subjects in which he is interested. Such a facility could be provided instead of or in addition to the 'push' news service described above.

CLAIMS

1. A method of operating a packet network to transmit a plurality of packets to respective different subsets of possible recipients, said method comprising the steps of, for one or more portions of the network, assigning a common local multicast address to packets are destined for different subsets of possible recipients but which share a subset of forwarding nodes within said portion.

2. A method according to claim 1 wherein said identification step involves:

extracting, from the payload of each packet, destination data identifying the subset of possible recipients to which the packet is to be sent; and

deriving an associated local multicast address on the basis of said destination data; and

transmitting said packet to said associated local multicast address.

3. A method according to claim 2 wherein said destination data comprises subject identifying data.

ABSTRACT

PACKET NETWORK

A packet network and method of operating the same is described. The network is arranged into a hierarchy, having one organisation-level portion (CS, P, R2) and three site-level portions (S1, S2, S3). By using forwarding computers and assigning multicast addresses in respect of each portion, the total number of multicast addresses required to send multicast messages to all possible subsets of potential recipients is reduced. This alleviates routers in the network of the burden associated with storing and processing a large amount of multicast addresses.

Figure (1)

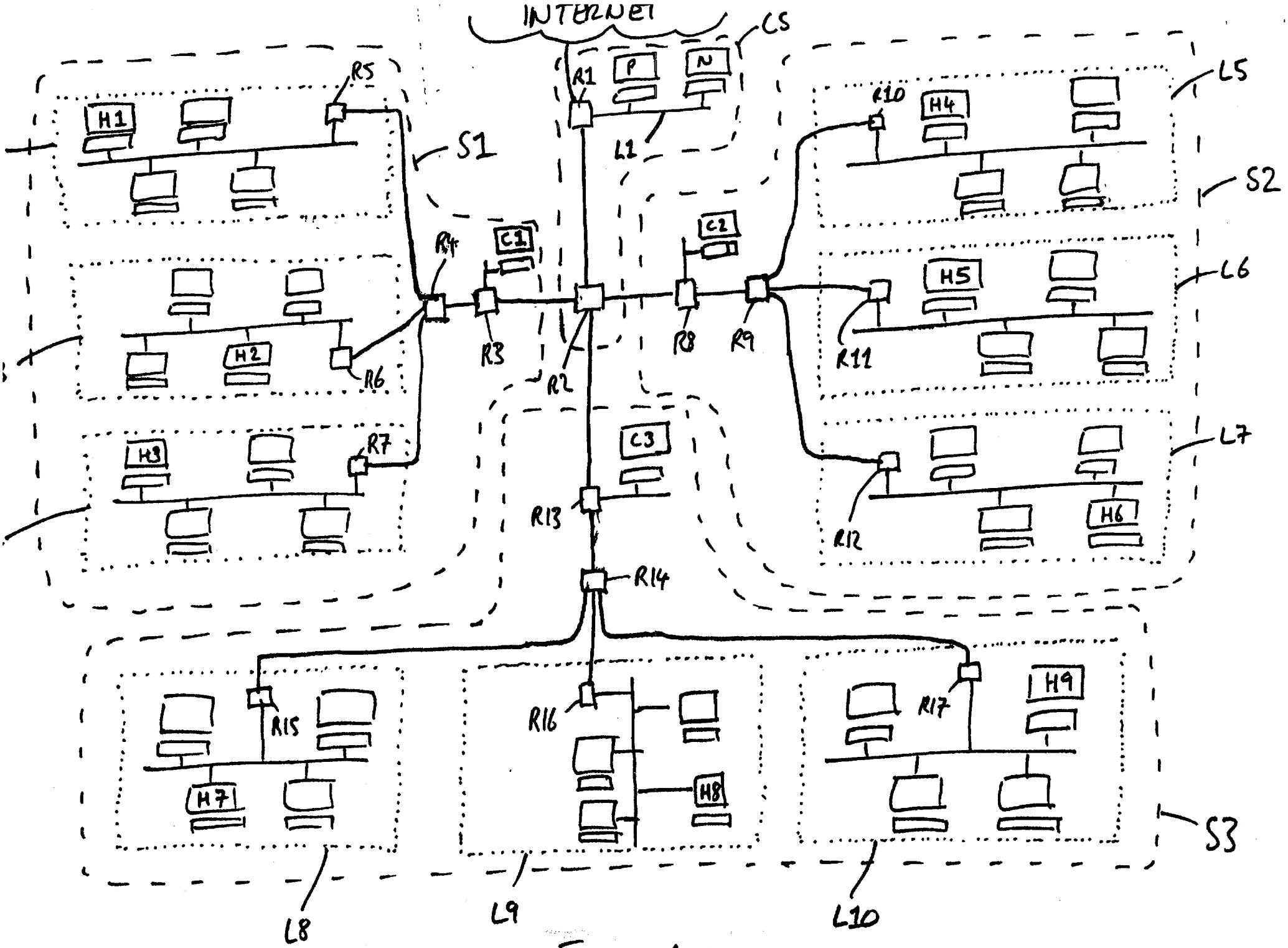
INTERNET
CS
R1
P
N
L1
L5
R10
H4
S2
S1
R5
H1
R4
C1
C2
L6
H5
R11
H2
R6
R3
R2
R8
R9
H3
R7
C3
L7
R13
R12
H6
R14
R15
R16
R17
H9
H7
H8
S3
L8
L9
L10

Figure 1

Members of Set (site level)	IP address
H1, H2, H3	230.81.56.1
H1, H2	230.81.56.2
H1, H3	230.81.56.3
H2, H3	230.81.56.4
H1	147.148.101.2
H2	147.148.102.2
H3	147.148.103.2

Figure 2A

Members of Set	IP address
C1, C2, C3	229:274:1:27
C1, C2	229:274:1:28
C1, C3	229:274:1:29
C2, C3	229:274:1:30
C1	147.144.101.2
C2	147.144.102.2
C3	147.144.103.2

Figure 2B

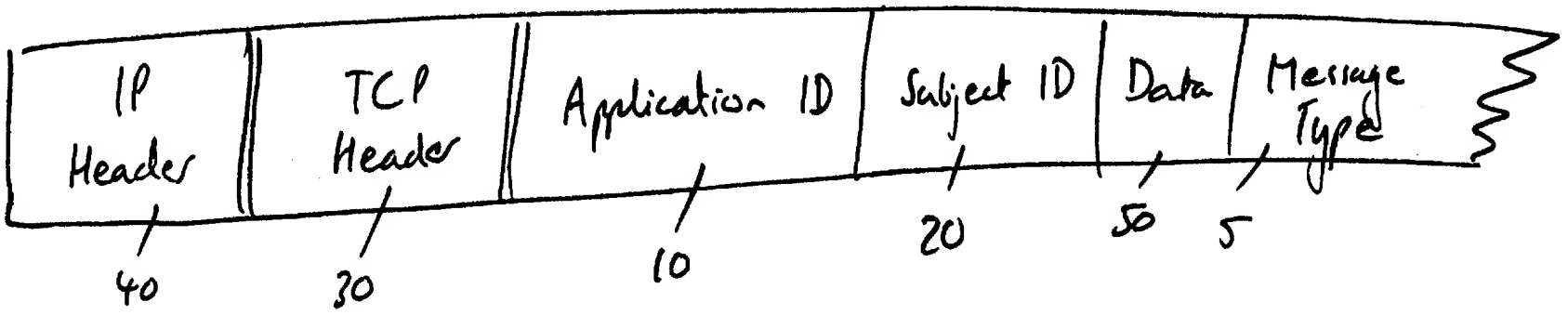
IP Header
TCP Header
Application ID
Subject ID
Data
Message Type
40
30
10
20
50
5

Figure 3

Members of Set (site-level)	IP address	Subjects subscribed to
H1, H2, H3	230.81.56.1	none
H1, H2	230.81.56.2	none
H1, H3	230.81.56.3	none
H2, H3	230.81.56.4	none
H1	147.148.101.2	pool
H2	147.148.102.2	none
H3	147.148.103.2	none

Figure 4A

Members of Set (organisation-level)	IP address	Subjects subscribed to
C1, C2, C3	229:274:1:27	none
C1, C2	229:274:1:28	none
C1, C3	229:274:1:29	none
C2, C3	229:274:1:30	none
C1	147.144.101.2	pool
C2	147.144.102.2	none
C3	147.144.103.2	none

Figure 4B

Members of Set (site-level)	IP address	Subjects subscribed to
H1, H2, H3	230.81.56.1	none
H1, H2	230.81.56.2	pool
H1, H3	230.81.56.3	none
H2, H3	230.81.56.4	none
H1	147.148.101.2	none
H2	147.148.102.2	none
H3	147.148.103.2	none

Figure 4C

Members of Set (site-level)	IP address	Subjects subscribed to
H1, H2, H3	230.81.56.1	pool
H1, H2	230.81.56.2	skiing
H1, H3	230.81.56.3	none
H2, H3	230.81.56.4	weight-lifting; wrestling
H1	147.148.101.2	cycling; golf
H2	147.148.102.2	rugby
H3	147.148.103.2	hockey

Figure 5A

Members of Set (site-level)	IP address	Subjects subscribed to
H4, H5, H6	230.81.58.1	swimming
H4, H5	230.81.58.2	fishing, pool
H4, H6	230.81.58.3	golf
H5, H6	230.81.58.4	none
H4	147.150.101.2	tennis
H5	147.150.102.2	basketball
H6	147.150.103.2	hockey

Figure 5B